Segmentation of Identical and Simultaneously Played Traditional Music Instruments using Adaptive

Yoyon K. Suprapto[1], Mauridhi H. Purnomo[1], Mochamad Hariadi[1]

Abstract — **Nowadays, mining of the musical ensemble has become very crucial since the information inside a musical ensemble is required by any musical contents services. In this research, we introduce Gamelan as one of the Indonesian traditional music instruments as our research objective. To indicate the changes of Gamelan features (i.e. tempo also the hammer struck styles) the segmentation of Gamelan music instruments is required as the music tagging tools. Adaptive LMS is employed for segmenting identical instruments that are played in the concurrent fashion. The target is to find how many instruments are played at the same time or separated by very short time (≤ 1 ms). The experiment results demonstrate robust detection with 0.02 ms accuracy for segmenting identical and simultaneously played Gamelan instruments. These results are employed for indicating the changes of Gamelan features, such as tempo also the hammer struck styles.**

Keywords — **component, Adaptive LMS, Gamelan features, Music tagging.**

I. INTRODUCTION

Nowadays digital media is developed from the most sophisticated music instruments until the traditional ones. Mining of the musical ensemble has become very crucial since the information inside a musical ensemble is required by any musical contents services. This is due to today concerns information overload, when the Internet is becoming the platform, the browser becomes the operating system and the applications become services [1].

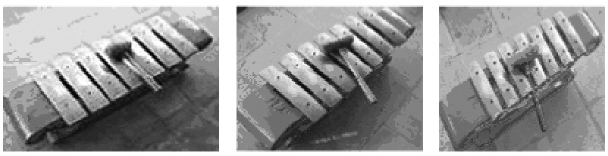

Fig. 1. Balungan, Demung, Saron and Peking.

Approaches to making the audio search space usable include text mining of music articles for semantic tags, hand tagging, query-by-humming, and lower level signal analysis and genre classification schema. More recently, researchers have turned to using human computation as a method to retrieve information from audio files.

In this paper, we introduce a technique of music tagging for an Indonesian Javanese traditional music instrument called *Gamelan*. A Gamelan set consists of several groups of different instruments. One of the groups is called *Balungan*.

Three instruments namely Demung, Saron and Peking [2] construct the Balungan group. Fig. 1 shows those instruments with their hammers. These hammers are employed to play these instruments by striking the blades. The type of Gamelan styles depend on how many identical instruments are played simultaneously along with the density among them. These will lead to indicate the changes of Gamelan features. Therefore, music tagging for Gamelan is very important to be applied in a Gamelan ensemble.

Gamelan ensemble used to be played with *Wayang* performance (Javanese Shadow Puppet Fig. 2) [3]. Therefore, the rhythms, themes and music styles of Gamelan depend on the story of Wayang performance. For instance, the audiences of Gamelan ensemble and Wayang performance could follow the story of Wayang performance by comprehending the rhythms, themes and music styles of Gamelan. This could be done by notifying how many identical instruments are played simultaneously [2]. Addressing this, music tagging for Gamelan requires the segmentation of these instruments with the density among them.

Fig. 2. Wayang performance

This paper describes a technique to segment identical and simultaneously played (e.g. separated by very short time (≤ 1 ms)) Gamelan music instruments. In [4] and [5], the polyphonic music instrument like piano, flute and guitar are segmented by employing a synthetic music instrument like MIDI. Each instrument that is played with different music notations can be segmented.

Nevertheless, the experiments do not employ identical and simultaneously played music instruments. In our

[1]Yoyon K. Suprapto, Mauridhi H. Purnomo, and Mochamad Hariadi are with Department of Electrical Engineering, Faculty of Industrial Technology, Institut Teknologi Sepuluh Nopember, Surabaya, 60111, Indonesia. E-mail: yoyonsuprapto@ee.its.ac.id, hery@ee.its.ac.id, mochar@ee.its.ac.id.

previous research [6] and [7], the sound of Saron instruments are extracted from a Gamelan ensemble using Inverse Filter [8]. Anyhow, the amount of identical instruments and the density among them are ignored. Regarding this requirement, we propose to employ adaptive LMS for segmenting identical instruments, which are played in simultaneous fashioned.

The structures of our paper are as follows: 2nd section describes the acoustic music instruments and its characteristics followed by adaptive LMS. 3rd section explain the design of the propose technique. The 4th section is experiments and 5th section concludes the paper and the road ahead.

II. Fundamental Theory

A. Balungan Signal

The *Balungan* group is constructed by three instruments namely *Demung, Saron and Peking* [2]. Each *Balungan* instrument has one octave, so the other octave is played by another instrument. *Demung* has the lowest octave, the range frequency is about 200 – 500 Hz, *Saron* has moderate octave, it is about 500 – 1000 Hz, and *Peking* has the highest octave, it is about 1000 - 2000 Hz [9]. A *Balungan* instrument is constructed by several *Gamelan blades*. Each blade represents a music notation.

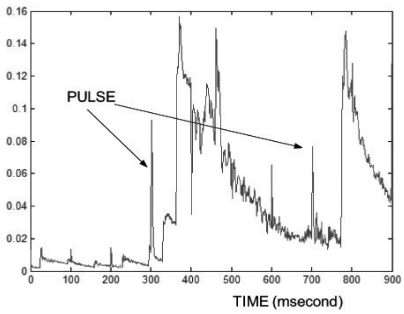

Fig. 3. Envelope of Balungan signal

If a Gamelan blade is struck by a hammer, the blade will vibrate and generate a sound. The sound will be absorbed within short period. In signal processing, the envelope of signal [8] is divided into three areas: attack, decay and sustain. During attack time, the amplitude grows up very fast and during decay time, the signal swings irregularly. The amplitude and the frequency density are uncontrolled. In decay period time, the amplitude goes smaller exponentially and the frequency becomes stable. Based on its character, the beginning of the signal can be detected by Amplitude steepest descent and uncontrolled Amplitude and Frequency density.

In Gamelan ensemble, we have some grammatical features and signal features. Grammatical features are customs how to play music like tempo, notations. Another feature is signal feature. Signal feature is signal characteristic like timbre, amplitude. During the music is played, both features are very often changed by a conductor. The tempo can be changed faster after the music is played by several notations, and than back to slower tempo. In additions, the technique to play an instrument is varied by another method likewise a music notation is struck twice, etc [2].

Each Gamelan instrument has a specific hammer. This hammer is employed to play these instruments by striking the blades. The sound character is influenced by the player temper. The hardness of blade struck and the blade struck style will influence the sound manners. In this research, we evaluate a *Balungan* attack time from several *Balungan* instruments. We use 200 data samples.

If an instrument is played individually, the beginning of the signal can be detected by Amplitude and frequency density. In Gamelan ensemble, the fastest tempo is 300 ms ([2]). So, if we play a number of instruments simultaneously, we have to extract each instrument sound, and then remove its background noise. We segment a notation sound from a Gamelan ensemble using adaptive LMS filter. Finally, we obtain the signal envelope described in Fig. 3.

Note: Some of data cannot be considered as signal but pulse (Fig. 4).

Definitions: Signal during time interval (is about 152 ms) and the amplitude difference more than 0.05 mV. Pulse time interval is less than 50 ms and the amplitude difference is greater than 0.05 mV.

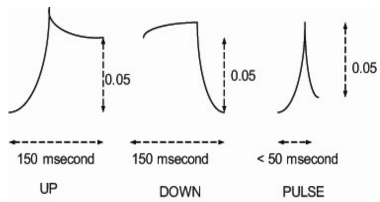

Fig. 4. Signal and Pulse

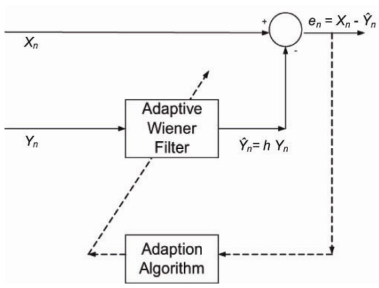

Fig. 5. Adaptive LMS

Separating two or more signals is a very important to find signal features, but separating two or more identical music instruments which is played simultaneously is an extraordinary thing [4] because they can be added each other, or can also be subtracted each other. It depends on the phase of the signals. If their phases are the same, both signals are added, otherwise they are subtracted each other.

B. Adaptive LMS Filter

Adaptive LMS filter is different from fixed point filter [8]. We design an adaptive linear system that responds to changes in its environment as it is operating. Linear networks that are adjusted at each time step based on new input and target vectors can find weights and biases that minimize the network's sum-squared error for recent input and target vectors. Networks of this sort are often used in error cancellation, signal processing, and control systems [10].

$$\hat{Y}_n = \sum_{m=0}^{M} h_n Y_{n-m} \quad (1)$$

$$e_n = x_n - \hat{Y}_n \quad (2)$$

$$\xi = E\|e^2\| = minimum \quad (3)$$

The value of eq. (3) is the minimum, when

$$\frac{\partial\zeta(\mathrm{h})}{\partial(h)} = 0 \quad (4)$$

$$\begin{aligned} h(n+1) &= h(n) + \Delta h(n) \\ &= h(n) + 2\mu e n y_n \end{aligned} \quad (5)$$

Where X_n is input signal, Y_n is reference signal, *h(n)* is filter weight, μ is step size, en is error signal, and $Y\hat{}_n$ is an output signal.

III. System Design

The algorithm of LMS [10] [11] are

1. At time n, the filter weight *h(n)* is available
2. Compute the filter output $\hat{Y}_n = \sum_{m=0}^{M} h_n Y_{n-m}$
3. Compute the estimation error $e_n = x_n - \hat{Y}_n$
4. Compute the next filter weight $h(n+1) = h(n) + 2\mu e n y_n$
5. Go to next time instant $n = n + 1$

The purpose of this technique is to evaluate filter weight to obtain minimum e_n If the input signal is a Gamelan ensemble, and the output signal is an extracted signal $\hat{Y}_n$. By using many Adaptive LMS iterations, the extracted signal can eliminate the background noise.

The $\hat{Y}_n$, extracted signal, does not generate information how many identical instruments are played simultaneously, separated by very short time interval ($\leq$ 0.1 ms), but Filter Weight *h(n)* can do it. Filter coefficient can detect how many identical music instruments are played in the same time and same music notation. These will lead to indicate the changes of Gamelan features, such as tempo, a hammer struck style, etc. Therefore, music tagging for Gamelan can be applied in a Gamelan ensemble.

IV. Experiment

A. Extracting a music instrument use Adaptive. LMS filter

Employing adaptive LMS we segment a music instrument from a Gamelan ensemble. Each music notation signal is separated from the ensemble and creates an output signal $\hat{Y}_n$. It represents a music notation, e.g. *Saron* 3 704 HZ), sound. Using amplitude detection, we are able to detect when a music notation played. (Fig. 8). Finally, we will generate the Gamelan features, such as a Gamelan notation and tempo.

What happen if amount of identical instruments among them are considered. We should detect how many identical instruments are played simultaneously along with the density among them. Gamelan instruments are played manually, therefore the players cannot play the music with extremely in the same time, but their identical sounds are separated by very short time interval $\leq$ 0.1 ms. The sound character is also influenced by the player temper. The hardness of hammer struck and the hammer struck style will influence the sound manners.

We employ adaptive LMS for segmenting identical instruments, which are played in simultaneous fashioned. The identical signals are used as input signals, X1 and X1' are *Saron* 3 notation (704 Hz) Signal. On the other hand we use another signal, *Y* is Saron3 notation too (704 Hz), as a reference signal. In this experiment, the filter length is 4000, μ step size is 0.001.After 60 times iteration, we can detect both signal, X1 and X1' are separated by 1700*thsampling*. The sampling rate is 48000, so 1700*thsampling* is equal with 1700*th/*48000 = 12 ms, by certain _t = 12 ms. (Fig. 9). We observe several experiments, which are described in Table 1.

Note: S1 is Saron1 (528 Hz), S2 is Saron2 (610 Hz), S3 is Saron3 (704 Hz), S5 is Saron5 (797 Hz), S6 is Saron6 (926 Hz).

B. Music Tagging.

In Gamelan ensemble, we have some grammatical features and signal features. Grammatical features talks about how to play the music such as tempo, music notations, kind of music, kind of instrument is played, and signal feature is signal characteristic like timbre, amplitude of signal. During the music is played, both music features, grammatical and signal, are very often changed by a conductor. The tempo can be changed faster after the music is played several notations, and back to slower tempo. In additions, how to play an instrument is sometimes varied by another method likewise a music notation is struck twice or with the other regulations, etc.

The type of Gamelan styles are depended on how many identical instruments are played simultaneously along with the density among them. These will lead to indicate the changes of Gamelan features. Therefore, music tagging for Gamelan is very important to be applied in a Gamelan ensemble, which each tagging represents a command. This code can be used as a command assignment, which represents the increasing or decreasing tempo, or changing a hammer struck technique. This code can also be used as a command assignment, which represents the hardness of hammer struck, and the hammer struck style.

The each tagging can be a combination ASCII code. For example, we insert a music code '11101101' in a Gamelan ensemble as a command assignment. Each '1', data bit signal, is represented by a *Saron* 3 sound (704 Hz), and '0', data bit signal, is represented by no sound (0 Hz). Each data bit is separated by very short time *0.1 ms* time interval that can be seen in Fig. 10. After 60 times iteration, we obtain the filter coefficient which can be seen in Fig. 11.

V. Conclusion

In this research, we employ adaptive LMS for segmenting identical Gamelan instruments called *Saron*, which are played in the concurrent fashion. Our target is to find how many *Sarons* are played at the same time, and find its density by segmentation. The Experiment results demonstrate the segmentation of simultaneously played *Sarons* with 0.02 ms in accuracy.

These will lead to the capability to indicate the changes of Gamelan features, such as tempo and a hammer struck style. These results would be useful for music tagging of Gamelan instruments. For the road ahead, adaptive LMS that we have developed could be made more accurate for different instruments of Gamelan.

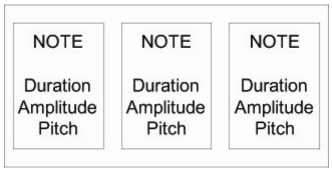

Fig. 7. A music tag sample

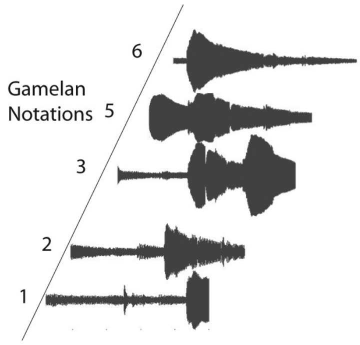

Fig. 8. Music notation separation

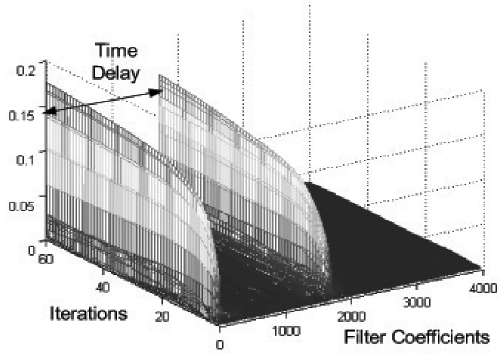

Fig. 9. Identical instrument detection

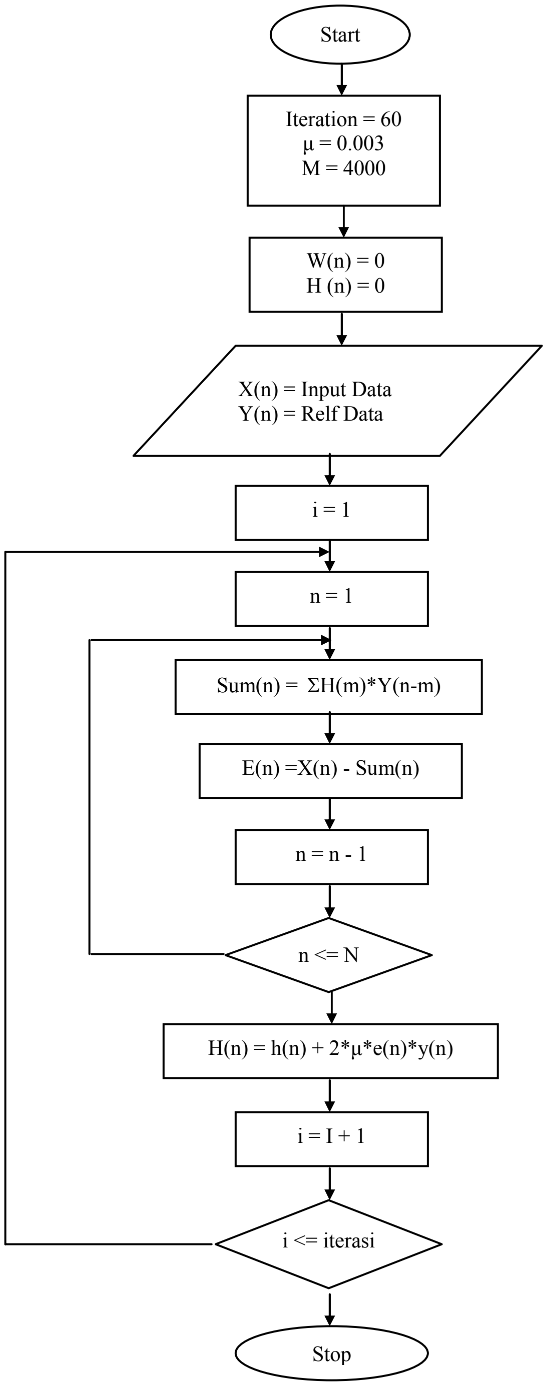

Fig. 6. Algorithm of adaptive LMS

TABLE 1
LMS ADAPTIVE EXPERIMENTS

Input Signal	Interval	Reff	Detected
*S*3 + *S*31	4 ms	S3	Yes
*S*3 + *S*31	12 ms	S3	Yes
*S*3 + *S*31	25 ms	S3	Yes
*S*3 +*S*31 + *S*1 +*S*5	10 ms	S3	Yes
*S*3 +*S*31 + *S*1 +*S*5	10 ms	S1	Yes
*S*3 + *S*31 + ... + *S*310	0.5 ms	S3	Yes
*S*3 + *S*31 + ... + *S*310	0.2 ms	S3	Yes

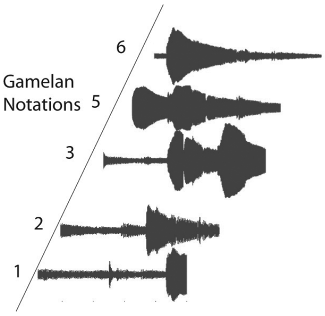

Fig. 10. Binary command code

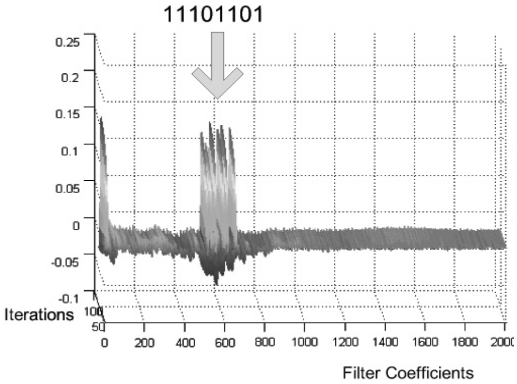

Fig. 11. A music tag

REFERENCES

[1] K. Glinos, 1997, "Information Access and Interfaces," *Presentation for the Last Call for the 4th Framework*, 29 september, Brussels."

[2] Sumarsam, 1992, "*Cultural Interaction and Musical Development in Central Java*," The University of Chicago Press.

[3] E. Salleh, 2008, "Wayang kulit," *National Library Board Singapore*, Dec.

[4] K.K.T.O. Tetsuro Kitahara, Masataka Goto and H.G. Okuno, 2007, "Instrument Identification in Polyphonic Music: Pitch Weighting to Minimize Influence of Sound Overlaps," *EURASIP Journal on Advance in Signal Processing*,vol.

[5] S.S. Haruto Takeda, Takuya Nishimoto, 2003, "Automatic Rhythm Transcription from Multiphonic Midi Signals,"*//citeseer.ist.psu .edu/ 652819.html*.

[6] M.H.P. Yoyon K Suprapto, Moch Hariadi, 2008, "A Gamelan Sound Segmentation using Inverse Filter," *The 6th Kumamoto University Forum,* Nov.

[7] M.H.P. Yoyon K Suprapto, Mochamad Hariadi, 2008, "Gamelan Sound Detection using Cross Correlation," Seminar on Intelligent Technology and Its Applications, 2008. J.O.S. III, "Physical audio signal processing," *Center for Computer Research in Music and Acoustics (CCRMA)*, Stanford University, Dec.

[8] B.L. Dhanda, 2003, "Perancangan dan Pembuatan Perangkat Lunak untuk Analisa dan Sintesa Bunyi Gamelan pada Komputer," *Thesis in Electrical Engineering Department ITS*.

[9] S.J. Orfanidis, 1988, "Optimum Signal Processing," *acMillan Publishing Company*, New York, vol.2nd edition.

[10] S. Haykin, 1996, "Adaptive Filter Theory," *Prentice Hall*, ol.3rd edition.